Privacy Region Protection for H.264/AVC with Enhanced Scrambling Effect and a Low Bitrate Overhead

Yongsheng Wang, Máire O'Neill, Fatih Kurugollu, and Elizabeth O'Sullivan

Centre for Secure Information Technologies
ECIT, Queen's University, Belfast, UK

Abstract

While video surveillance systems have become ubiquitous in our daily lives, they have introduced concerns over privacy invasion. Recent research to address these privacy issues includes a focus on privacy region protection, whereby existing video scrambling techniques are applied to specific regions of interest (ROI) in a video while the background is left unchanged. Most previous work in this area has only focussed on encrypting the sign bits of nonzero coefficients in the privacy region, which produces a relatively weak scrambling effect. In this paper, to enhance the scrambling effect for privacy protection, it is proposed to encrypt the intra prediction modes (IPM) in addition to the sign bits of nonzero coefficients (SNC) within the privacy region. A major issue with utilising encryption of IPM is that drift error is introduced outside the region of interest. Therefore, a re-encoding method, which is integrated with the encryption of IPM, is also proposed to remove drift error. Compared with a previous technique that uses encryption of IPM, the proposed re-encoding method offers savings in the bitrate overhead while completely removing the drift error. Experimental results and analysis based on H.264/AVC were carried out to verify the effectiveness of the proposed methods. In addition, a spiral binary mask mechanism is proposed that can reduce the bitrate overhead incurred by flagging the position of the privacy region. A definition of the syntax structure for the spiral binary mask is given. As a result of the proposed techniques, the privacy regions in a video sequence can be effectively protected by the enhanced scrambling effect with no drift error and a lower bitrate overhead.

Keywords: Privacy region protection, selective encryption, ROI, drift error, scrambling effect, bitrate overhead, binary mask, sign bits, nonzero coefficients, intra prediction mode, H.264/AVC.

1. Introduction

To defeat the increasing terrorism and criminality in our society, video surveillance systems are now widely deployed and can be found in shopping malls, streets, schools, banks, public transportation stations, private properties, and so on. The enhanced security sense provided by this wide deployment of video surveillance systems is welcomed by most people. However, the invasion of privacy and the possible abuse of video surveillance systems has drawn serious concern [1, 2, 3, 4, 5]. Even in public spaces, people still have a right to privacy. It has been reported [2] that "many European states explicitly acknowledge that CCTV surveillance in public space creates a conflict with the right to privacy." People have the right to demand that their privacy is properly protected in surveillance video using advanced processing techniques.

Email address: ywang26@qub.ac.uk, m.oneill@ecit.qub.ac.uk, f.kurugollu@qub.ac.uk, e.osullivan@qub.ac.uk (Yongsheng Wang, Máire O'Neill, Fatih Kurugollu, and Elizabeth O'Sullivan)

To address this, significant research has been conducted into protecting specific regions of interest (ROI) in video surveillance systems, such as people's faces in each frame of the video, or moving objects. The principle idea is to scramble or obscure the regions of interest while leaving the background in the clear for normal monitoring.

Since H.264/AVC is reported to have a better performance compared to other previous video compression standards, it has been widely adopted in many industrial applications [6, 7, 8]. One of its main applications is in video surveillance systems. The research in this paper focuses on privacy region protection based on H.264/AVC. Two main issues are discussed: one is on how to enhance the scrambling effect of the privacy region without causing drift error; the other is how to more effectively flag the position of the privacy region. The precise detection of the privacy region, which involves motion detection and object tracking techniques, is outside the scope of this work and it is assumed to be known.

Previous work in this field includes two kinds of methodologies to protect privacy in video surveillance systems. One adopts computer vision or video processing techniques to obscure the regions of interest [1] [9, 10, 11, 12], also known as non-scrambling based techniques. The second one is to scramble the regions of interest by video encryption techniques [13 - 25], also known as scrambling based techniques. These will be reviewed in detail in Section 2. It is believed that the methodology based on video encryption is more reliable and secure since video encryption has been investigated very well in the last decade [26, 27, 28, 29, 30] and it is based on cryptographic algorithms. Since a scrambled video still retains the information to regain the original frame with the decryption key, the main advantage of the scrambling based techniques is their capability to recover the original frame for later reference. Compared with non-scrambling techniques, scrambling based techniques can provide a flexible way to manage privacy-related privileges. The encrypted video is distributed and only an authorized person with knowledge of the correct key can descramble the privacy region, otherwise it remains scrambled while the background remains clear for normal monitoring. Privacy-related privileges rely on key management. Most previous work only utilized the encryption of the sign bits of nonzero coefficients (SNC) in the privacy region. However, as shown in Section 4.1, on its own this technique does not provide a very effective scrambling effect. To enhance the scrambling effect, in this paper, it is proposed to also encrypt the intra prediction modes (IPM) in the privacy region [21]. Directly applying scrambling techniques for privacy region protection in H.264/AVC causes a drift error in the non-privacy region, which is undesirable for monitoring applications. Very recently, Peng *et al.* [22] proposed to also encrypt the intra prediction modes and motion vectors in the privacy region in addition to the sign bits of nonzero coefficients. However, as they indicated in their paper, they did not address the drift error in the non-privacy region caused by encryption of the ROI. In this work, a new method using a re-encoding process is proposed to eliminate the drift error when encrypting the intra prediction modes in the privacy region [21]. This proposed method can completely remove the drift error in the non-privacy region while the scrambling effect in the privacy region is effectively enhanced. Compared with previous work that involves the encryption of IPM, the proposed method can offer savings in the bitrate overhead. To more efficiently indicate the position of the privacy region, a spiral binary mask mechanism is also proposed. This improves upon previous work involving an adaptive binary mask mechanism [31], which used binary arithmetic coding to compress the binary mask of each frame. In summary, the main contributions of this paper include:

1) a detailed review of previous work in privacy protection in video surveillance systems.

2) the combination of encryption of IPM and SNC to enhance the scrambling effect in the privacy region with experimental results for higher resolution videos.

3) a re-encoding method to remove the drift error caused by the encryption of IPM with experimental results for higher resolution videos.

4) a spiral binary mask to more efficiently indicate the position of the privacy region.

The rest of this paper is arranged as follows. Previous work on privacy protection in video surveillance systems is reviewed in Section 2. In Section 3, the existing work on how to prevent the drift error incurred by scrambling the privacy region is discussed. Section 4 illustrates the weak scrambling effect that results from only encrypting the sign bits of nonzero coefficients in the privacy region, and also presents the proposed

re-encoding method, which can enable the encryption of intra prediction modes in the privacy region to enhance the scrambling effect without drift error. In Section 5, the issue of how to efficiently flag the position of the privacy region is discussed and a novel spiral binary mask mechanism is presented to reduce the bitrate overhead. Experimental results and analysis are also given in Section 4 and 5 to illustrate the effectiveness of the proposed methods. Finally, conclusions are drawn in Section 6.

2. Previous work on privacy protection in video surveillance systems

Previous work on privacy protection in video surveillance systems is categorized into two kinds: non-scrambling based techniques and scrambling based techniques. The work in this paper belongs to the latter, which is regarded as being more reliable and secure.

2.1. Non-scrambling Based Techniques

Senior *et al.* [1] proposed a privacy-preserving surveillance system to manage access to different versions of video-derived data according to access-control lists. In their work, they concentrate on the data processing policy of a privacy-preserving system and do not specify the detailed techniques on how to encrypt or scramble the privacy region. Another problem in their system is that there is too much rendering of videos for different access levels which may not be necessary but increases the computational complexity and affects the real-time performance. Newton *et al.* [9] proposed a technique to defeat facial recognition software by "de-identifying" a face image, however the face image after processing was still sufficiently visible for human eyes to identify, which indicated that facial recognition software could be further improved against their technique. Zhang *et al.* [10] proposed to hide the data in the privacy region in the frame itself as a watermark in H.263. But the increase in bitrate is overwhelming and the decreased perceptual quality of the non-privacy region is undesirable. In the work of Martínez-Ponte *et al.* [11], based on JPEG 2000, the data in the privacy region was separated as a single quality layer and without this layer, the privacy region was invisible or blurred. Yu *et al.* [12] first discussed the definition of privacy from a psychological perspective and proposed to replace the privacy region by different visual objects, such as silhouette, border and so on. But in the resulting video, the privacy region is irrecoverable. Thus, they provided another recoverable method to mask the face according to a training model. However, in their research they fail to consider the impact on the video compression performance. All of the previous work discussed above focused on obscuring the privacy region to some extent by non-scrambling techniques. The data in the privacy region was not protected by any cryptographic technique.

2.2. Scrambling Based Techniques

Recently, there has been research [13 - 25] carried out to protect the private content in a video by applying selective encryption to the privacy region only. Selective encryption techniques are well studied [26, 27, 28, 29, 30]. Generally they are based on some form of cryptographic algorithm, thus the security is more reliable than non-scrambling based techniques. The encrypted video can be decoded as normal, but without decryption, the privacy region remains scrambled while the remainder of each frame, the non-privacy region, is left in the clear. If requested in future, the encrypted privacy region in the video can be decrypted to show the original. Based on JPEG 2000 and MPEG-4, Dufaux *et al.* [13] [14] first proposed to scramble the sign bits of nonzero AC transform coefficients of blocks in the privacy region. An MB-type (here MB is short for macroblock) decision mechanism is employed to hinder the drift error caused by inter prediction, which means that the macroblock in the current frame, collocated with a MB in the privacy region of reference frames, is always intra coded. The authors extended their technique to the codestream-domain of MPEG-4 [15]. The compressed video stream is firstly parsed and then the corresponding sign bits of the AC transform coefficients are encrypted as described above. In [16], their technique was further extended to H.264/AVC, which was the first scheme for privacy region protection based on H.264/AVC. The Flexible Macroblock Ordering (FMO) mechanism of H.264/AVC [6, 7, 32] is utilized to indicate which macroblocks belong to the privacy region and prevent drift error due to intra prediction from the privacy region to the non-privacy region. The MB-type decision mechanism is also adopted to eliminate drift error due to inter

prediction. Baaziz *et al.* [17] built a complete automated video surveillance system by combining the method of Dufaux *et al.* [13] [14] with motion detection and watermarking. In [18], a permutation based encryption method was used to scramble the privacy region in the spatial domain before compression but the resulting increase of the bitrate was very high, around 23% on average according to their experimental results for H.264. Based on Motion JPEG 2000, Martin *et al.* [19] utilized a stream cipher to obscure visual objects by encrypting some specific bits in the compressed bitstream, however, compared with H.264/AVC, Motion JPEG 2000 is not commonly used in video surveillance systems because of the compression performance. Sohn *et al.* [20] pointed out that only encrypting the sign bits of nonzero AC coefficient was not secure against the error concealment attack and proposed to randomly flip the sign bits of both the nonzero DC and AC coefficients for H.264/SVC. In all of this previous research only the sign bits of nonzero coefficients (denoted as SNC) in the privacy region are encrypted. However, this produces a relatively weak scrambling effect, as shown in Section 4.1. To enhance the scrambling effect, we first proposed to also encrypt the intra prediction modes in addition to the sign bits of nonzero coefficients in the privacy region [21]. Very recently, utilising FMO and chaos, Peng *et al.* [22] also proposed to encrypt the intra prediction modes and motion vectors in the privacy region. However, as they indicated in their paper, the non-privacy region was affected by the encryption of the privacy region due to the drift error. Further more, using FMO can seriously affect the compression ratio. The serious impact on the compression performance is a common drawback for previous work using FMO to remove the drift error and flag the position of the privacy region in each frame [23] [24]. It is reported that in the worst case, the bitrate overhead using FMO exceeds 200% [23] [24]. In the work of [23] and [24], they proposed to remove the drift error using new techniques (reviewed in next section) and indicate the position of the privacy region using a binary mask. The compression ratio can be significantly improved compared with previous work in [16] with a bitrate overhead of less than 16%. Here, a binary mask means that one bit per MB is used to flag whether or not it belongs to the privacy region.

3. Review of the drift error for privacy region protection in H.264/AVC

In H.264/AVC, directly encrypting the intra prediction modes (IPM) [28] and/or the sign bits of nonzero coefficients (SNC) [16] [26] in the privacy region of a video frame will result in drift error in the non-privacy region due to intra and/or inter prediction from the privacy region. Two examples of the drift error when applying SNC in the privacy region are shown in Figure 1: (a) shows the drift error due to the intra prediction in the 15th frame of 'foreman'; (b) shows the drift error due to the inter prediction in the 40th frame of 'hall'. The drift error is indicated by the red ellipses. Thus, additional processing is required to prevent such drift error.

In this section, the previous work to remove the drift error caused by encrypting the SNC in the privacy region is first reviewed. Then, the two main methods (MRIP and SWRME) in this previous work are further explained in Section 3.2 and 3.3. MRIP is used to prevent the drift error caused by intra prediction when applying the SNC. SWRME can remove the drift error due to inter prediction. In the last section, an improved MRIP is examined, which can prevent the drift error caused by the intra prediction when encrypting the intra prediction modes.

3.1. Review of Existing Work on Removing Drift Error

Choi *et al.* [25] discussed the drift error that occurs when applying SNC to privacy region protection in H.264/AVC. In order to prevent drift error due to intra prediction, the intra prediction modes of blocks to the right and bottom of the privacy region are modified to exclude prediction from the privacy region. The same MB-type mechanism as that used by Dufaux *et al.* [14] [15] is adopted. In addition, the privacy region in reference frames can not be used to predict blocks in the non-privacy region and their motion vectors. However, Choi *et al.* did not evaluate the impact on the compression performance. Tong *et al.* [23] proposed similar solutions to avoid drift error when applying different selective encryption methods to privacy regions based on H.264/AVC. They proposed two techniques, Mode Restricted Intra Prediction (MRIP) for intra prediction and Search Window Restricted Motion Estimation (SWRME) for inter prediction, to remove drift error when encrypting the sign bits of nonzero transform coefficients. These are described later in this

(a) Drift error due to intra prediction

(b) Drift error due to inter prediction

Figure 1. Examples of the drift error, indicated in the red ellipses. It is better to view artifacts in the electronic copy.

section. They also proposed to exclude the use of the intra 4x4 prediction modes of blocks to the right and bottom of the privacy region when encrypting the intra prediction modes in the privacy region. In addition, a binary mask (one bit per MB) is employed to indicate the privacy region without using the slice group separation technique [7] [32] as used in FMO (Flexible Macroblock Order). Dai *et al.* [24] extended the work by Tong *et al.* [23] by introducing the Boundary Strength Restricted Deblocking Filter (BSRDF) to further remove the drift error (In H.264/AVC, the deblocking filter is optional and the incurred drift error is slight. In this paper, for simplicity, the deblocking filter is disabled.). Compared with the work by Dufaux *et al.* [16] based on FMO, the methods by Tong *et al.* [23] and Dai *et al.* [24] have significantly improved the bitrate overhead incurred due to the privacy region protection while still preventing drift error. The main reason for the significant reduction in the bitrate overhead is because the FMO is disabled and instead a binary mask to indicate the position of the privacy region and new techniques to remove the drift error are used. However, this previous work mainly focuses on encrypting only the sign bits of nonzero transform coefficients. The scrambling effect when only encrypting the sign bits is relatively weak and can not conceal all the details of the privacy region.

3.2. Mode Restricted Intra Prediction (MRIP)

MRIP was proposed separately by Tong *et al.* [23] and Choi *et al.* [25]. In H.264/AVC, an intra encoded macroblock may be encoded into the intra 4x4 prediction mode or the intra 16x16 prediction mode. There are nine intra 4x4 prediction modes and four intra 16x16 prediction modes. The fundamental idea in the MRIP technique is to restrict the possible intra prediction modes for blocks around the boundary of the privacy region. As shown in Figure 2, for a block, C, in the non-privacy region, some of the nine intra 4x4 prediction modes are excluded. This exclusion applies only if any of the blocks, to the top (T), left (L), top-left (TL) and top-right (TR) of the current block, C, are located within the privacy region. The excluded intra 4x4 prediction modes are listed in Table 1. In particular, if both the top and left blocks of the current block, C, belong to the privacy region, all nine intra 4x4 prediction modes are excluded and the current block is coded in IPCM (Intra Pulse Code Modulation) [6] [7] which does not need any prediction from its adjacent blocks. This is also the case for intra 16x16 prediction. Table 2 lists excluded modes for the 16x16 block according to its neighbouring block position relative to the privacy region.

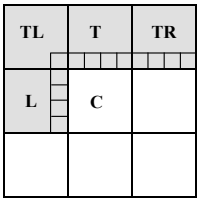


Figure 2. The adjacent blocks that may affect the current block by intra prediction.

Table 1. The excluded intra 4x4 prediction modes of a 4x4 block if its adjacent block is within the privacy region

Position	The excluded intra prediction modes
T	DC, Diagonal down left, Diagonal down right, Vertical left, Vertical right and Horizontal down
L	Horizontal, DC, Diagonal down right, Vertical right, Horizontal down and Horizontal up
TL	DC, Diagonal down right, Vertical right and Horizontal down
TR	Diagonal down left and Vertical right

3.3. Search Window Restricted Motion Estimation (SWRME)

To remove the drift error caused by inter prediction from the privacy region to the non-privacy region, this form of inter prediction can be forbidden, as shown in Figure 3. This is also referred to as search window

Table 2. The excluded intra prediction modes of a 16x16 block if its adjacent block is within the privacy region

Position	The excluded intra 16x16 prediction modes
T	Vertical, DC and Plane
L	Horizontal, DC and Plane
TL	Plane
TR	None

restricted motion estimation (SWRME) [24]. For example, in Figure 3, it is forbidden to use any block in the privacy region of the reference frame, such as B, to predict a block in the non-privacy region of the current frame, such as block A.

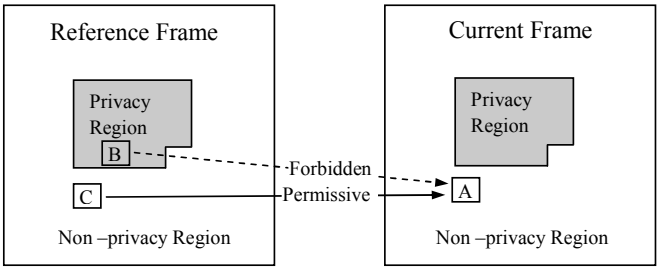


Figure 3. The inter prediction from the privacy region to the non-privacy region is forbidden to remove the drift error.

In addition, as stated by Tong *et al.* [23] and Dai *et al.* [24], it should also be forbidden for a block in the non-privacy region, but near the privacy region, to use the sub-pixel interpolation from the privacy region for motion estimation with half and quarter sample accuracy, to ensure that drift error can be removed. For example, as show in Figure 4, the half sample b would be interpolated from its 6 nearest pixels in a horizontal direction, according to equation (1). Since I and J are located in the privacy region and have been scrambled, the half sample b would be wrongly calculated to cause a drift error in the non-privacy region.

$$b = \text{round}(E - 5F + 20G + 20H - 5I + J)/32 \tag{1}$$

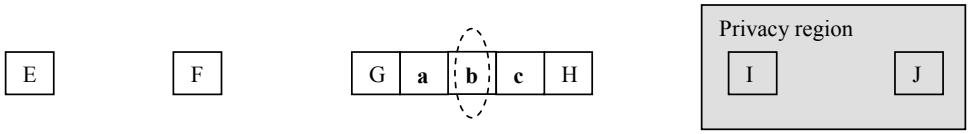


Figure 4. Half and quarter samples near the privacy region are interpolated from privacy region pixels.

3.4. Improved MRIP for Encryption of IPM

MRIP was designed to prevent the drift error caused by intra prediction when encrypting the sign bits of nonzero quantized DCT coefficients. Tong *et al.* [23] discussed the scrambling of the intra prediction modes (similar to the method by Ahn *et al.* [28]) in the privacy region and they also presented an improved MRIP to prevent the corresponding drift error. To encode the intra 4x4 prediction modes of the current block, C, its most probable intra prediction mode (denoted as $\text{IPM}_{probable}$ in Figure 5) will be the smaller of the two intra 4x4 prediction modes of its neighbouring top and left blocks, T and L. If either of these is not encoded by the intra 4x4 prediction mode, the most probable intra 4x4 mode is set to the DC prediction mode. If the best intra 4x4 prediction mode (denoted as IPM_{best} in Figure 5) for optimizing the distortion [7] equals the most probable one, a flag bit 'prev_intra4x4_pred_mode' is set with a value of '1'; otherwise, the flag bit is set to '0' and a 3-bit parameter 'rem_intra4x4_pred_mode' is followed. If the most probable intra 4x4

prediction mode is greater than the best mode, the value of 'rem_intra4x4_pred_mode' equals the best mode; otherwise, it is set to the best mode minus 1. Figure 5 explains this encoding process for intra prediction modes. The intra 4x4 prediction mode of the current block is coded according to the most probable mode which is predicted from the coded modes of the two neighbouring blocks (top and left). After encrypting the intra prediction modes of blocks in the privacy region, the intra prediction modes of blocks to the right and bottom of the privacy region are very likely to be decoded incorrectly. Tong *et al.* [23] proposed to exclude the intra 4x4 prediction mode for blocks if their left or top block is in the privacy region. In all other cases, Table 1 and 2 are still used.

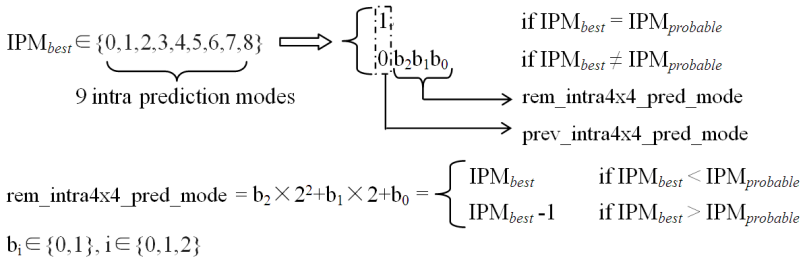


Figure 5. Explanation on how to encode the intra prediction mode.

4. Encrypting the intra prediction modes in the privacy region

This section first shows the weak scrambling effect when only applying encryption of the sign bits of nonzero quantized transform coefficients (SNC) for privacy region protection. Next, a proposal to enhance the scrambling effect by further encrypting the IPM in the privacy region is presented. To remove the drift error caused by encrypting the IPM, a re-encoding method is proposed in Section 4.3. Finally, experimental results are given to show the enhanced scrambling effect and the savings in the bitrate overhead incurred by the proposed privacy region protection methodology.

For encryption, a stream cipher is used in this research, the decryption of which is actually the same as its encryption. The encryption process simply involves XORing P_i with K_i to obtain C_i, as demonstrated in equation 2. Here, P_i indicates the i-th syntax element to be encrypted in the video stream; C_i is the corresponding cipher text; K_i is is a short string of random bits of the same length as P_i, which is generated by the stream cipher. For decryption, K_i is XORed with C_i to give P_i.

$$C_i = P_i \oplus K_i \tag{2}$$

4.1. The Weak Scrambling Effect of SNC

As mentioned in Section 1, most previous scrambling based methods for privacy region protection [13 - 25] only involve encryption of the SNC. However, on its own, this technique can not provide a strong scrambling effect. In Figure 6, the facial region in the 1st and 15th frames of 'foreman' in CIF (Common Intermediate Format) resolution is encrypted by SNC, which is encoded in CAVLC (Context Adaptive Variable Length Coding) under the baseline profile. It is clear that the scrambling effect is not sufficient to conceal the full details of the face. Thus, to effectively protect the privacy of sensitive information, it is necessary to enhance the resulting scrambling effect.

4.2. Combination of IPM and SNC in the Privacy Region

There are three main scrambling techniques that can be employed for H.264/AVC: encrypting the intra prediction modes, denoted as IPM [28]; encrypting the sign bits of the nonzero coefficients, denoted as SNC [26] [27]; encrypting the sign bits of motion vectors, denoted as EMV [27]. When encrypting the whole picture in each frame, the combination of these three techniques can provide a very strong scrambling effect [30] [33]. However, in most previous work on privacy region protection, only the sign bits of nonzero

coefficients are encrypted because using this technique it is easy to remove the drift error, as reviewed in Section 3. However, the resulting scrambling effect is relatively weak.

(a) The 1st frame

(b) the 15th frame

Figure 6. The relatively weak scrambling effect in the 1st and 15th frames of 'foreman' only with SNC for the privacy region protection [21]: the 1st frame is an I frame and the 15th frame is a P frame. Encoding setting: baseline profile (IPP..IPP..), CAVLC, QP=24 and intra period=5.

Since nonzero coefficients represent some details of the privacy region, they must be encrypted to conceal any sensitive information as carried out in previous work. As shown in previous work by the authors [30], of the three scrambling techniques for H.264/AVC, IPM is the most efficient and can achieve a better scrambling effect than EMV. When encrypting the intra prediction modes in the privacy region, the privacy region in I frames can be seriously scrambled because of the decoding error of intra prediction modes, and the privacy regions in P and B frames can also be effectively degraded because of the inter prediction from the scrambled privacy region in I frames. However, for privacy region protection, the authors believe that EMV is not necessary for the following reasons: The application of EMV only scrambles the motion inside the privacy region; however, the privacy region is very likely to be an object which will be moving in the same direction. Since only the privacy region is encrypted and the background is left clear, the movement of the privacy region as a whole object is still visible even if EMV is employed. Also, using EMV requires more

random bits for encryption in addition to complicated methods to remove the resulting drift error, both of which will increase the complexity of the system. Thus, to enhance the scrambling effect of the privacy region, the SNC and IPM methods are combined in this research. To remove the drift error caused by IPM in the privacy region, new techniques are needed. Therefore, in the next section, a re-encoding method is proposed, which can reduce the bitrate overhead compared with the technique (Improved MRIP) proposed by Tong *et al.* [23], which, to the best of our knowledge, is the only other research that looks at removing the drift error caused by the encryption of IPM.

4.3. Re-encoding IPMs around the Boundary

Generally, macroblocks those are encoded with the intra 4x4 prediction mode as the best mode include more texture than those macroblocks with intra 16x16 prediction mode or IPCM (Intra Pulse Code Modulation) as the best mode, and if encoded in other modes, the compression performance will be affected and the bitrate overhead will be caused. As indicated in Section 3.4, to avoid the drift error when encrypting the intra prediction modes in the privacy region, the non-best intra prediction mode for macroblocks around the boundary is likely to be used instead of the best intra prediction mode. Thus, a new mechanism [21] is proposed here to encrypt the intra prediction modes in the privacy region such that the best intra 4x4 prediction mode can still be used for the blocks around the boundary. Each of the intra 4x4 prediction modes of the blocks within the privacy region is encrypted by XORing with three random bits generated by a secure stream cipher. When encoding the intra 4x4 prediction mode of a block around the boundary, if any one of its adjacent top or left blocks is located within the privacy region, the most probable prediction mode should be calculated from the encrypted intra 4x4 prediction modes of the two neighbouring blocks. The two parameters, 'prev_intra4x4_pred_mode' and 'rem_intra4x4_pred_mode' are then sent in the normal way as specified in the H.264/AVC standard. An example of the proposed mechanism is show in Figure 7, in which the block above the current block has an intra 4x4 prediction mode of 5 and is within the privacy region, and the block to the left has an intra 4x4 prediction mode of 4 and does not belong to the privacy region. Three parameters (α, β, γ) are defined to represent the information for the encoding of each block. Here, θ corresponds to the best intra 4x4 prediction mode, α represents the corresponding most probable intra prediction mode, and the two parameters, β and γ, indicate 'prev_intra4x4_pred_mode' and 'rem_intra4x4_pred_mode', respectively. Given that the intra 4x4 prediction mode of the current block (which is located around the boundary of the privacy region) is 3, after encryption, the intra prediction mode of its top block is scrambled to 1 as shown in Figure 7(b). Without any special processing in the decoding procedure, the current block's intra prediction mode will be decoded incorrectly as 4 (Figure 7(c)). Furthermore, this error can propagate into subsequent blocks, which will potentially enhance the drift error. To solve this problem, after encrypting the top block's IPM, the intra prediction mode of the current block should be re-encoded again according to the encrypted intra prediction mode. This process is also explained by the pseudo code in Algorithm 1. As shown in Figure 7(d), the re-encoded intra prediction mode is decoded correctly according to the encrypted adjacent IPMs. As a result, the decoding error that typically happens when encrypting IPM is avoided. In the decryption procedure, the intra 4x4 prediction modes of blocks around the boundary should be decoded according to the adjacent encrypted IPMs as well.

4.4. The Enhanced Scrambling Effect

In order to evaluate the proposed methods, experiments were conducted based on the JM 17.2 [34] reference software. The stream cipher, 'Rabbit' [35], was used to generate the random bit sequence required during the scrambling process and five test sequences, 'foreman', 'road' and 'hall', in CIF resolution, 'crew' and 'BQMall' in 4CIF resolution (704×576 for 'crew' and 832×480 for 'BQMall'), were chosen for the evaluation. The first 30 frames of 'foreman', 'road', 'crew' and 'BQMall', and the second 30 frames of 'hall' from the 31st frame to the 60th frame, are encoded in the baseline profile and the main profile with CAVLC as the entropy coding method. In 'foreman', 'crew' and 'BQMall', the facial regions are defined as the privacy region, and in 'road', the moving car is regarded as the privacy region. In the first 30 frames of 'hall', the people moving in the scene, which form the privacy region, only appear briefly, thus the second 30 frames are used. The combination of IPM and SNC is used to evaluate the proposed re-encoding method

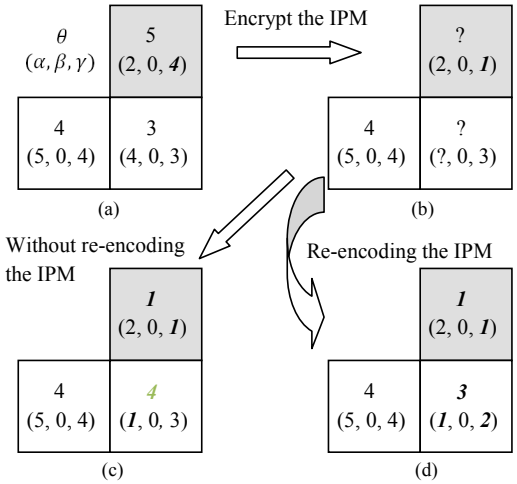


Figure 7. An example of re-encoding the intra 4x4 prediction mode of a block around the boundary of the privacy region [21].

Algorithm 1 Re-encoding IPMs around the boundary of the ROI

Step 1: If the current block is not intra-coded, go to Step 7.
Step 2: Encode the intra prediction mode (θ) of the current block as outlined in Figure 5.
Step 3: If the current MB is not in the ROI, go to Step 5.
Step 4: Encrypt the intra prediction mode by XORing with three random bits, then go to Step 8.
Step 5: If the current MB is at the boundary of the ROI,
Step 5.1: If the top block of the current block is in the ROI, re-calculate the θ_t of the top block
Step 5.2: If the left block of the current block is in the ROI, re-calculate the θ_l of the left block.
Step 5.3: According to the results above, update the $\alpha = \min(\theta_t, \theta_l)$ of the current block.
Step 5.4: Re-encode the θ of the current block as outlined in Figure 5.
Step 6: Go to Step 8.
Step 7: Encode the current block via inter prediction.
Step 8: Encode the next block and go back to Step 1.

(a) The 1st frame

(b) the 15th frame

Figure 8. The 1st and 15th frames of 'foreman' using IPM and SNC for privacy region protection [21]: the 1st frame is an I frame and the 15th frame is a P frame. Encoding setting: baseline profile (IPP..IPP..), CAVLC, QP=24 and intra period=5.

(a) The 15th frame of 'road' encrypted by SNC.

(b) The 15th frame of 'road' encrypted by IPM+SNC.

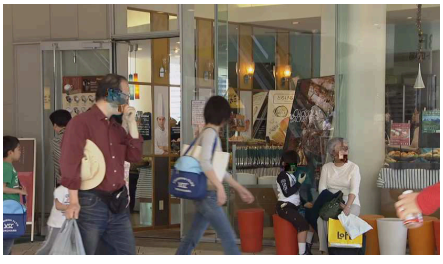
(c) The 40th frame of 'hall' encrypted by SNC.

(d) The 40th frame of 'hall' encrypted by IPM+SNC.

(e) The 15th frame of 'crew' encrypted by SNC.

(f) The 15th frame of 'crew' encrypted by IPM+SNC.

(g) The 15th frame of 'BQMall' encrypted by SNC.



(h) The 15th frame of 'BQMall' encrypted by IPM+SNC.

Figure 9. Example of the enhanced scrambling effect using IPM and SNC. All these example frames are P frames. Encoding setting: baseline profile (IPP..IPP..), CAVLC, QP=24 and intra period=5.

and a comparison with the work of Tong *et al.* [23] is provided. The privacy region is flagged on a MB basis (one bit per MB). Here, SWRME [23] [24] is adopted to prevent drift error for inter prediction.

As mentioned in Section 4.1, solely using SNC for privacy region protection results in a weak scrambling effect. Thus, it is better to apply IPM and SNC together to enhance the scrambling effect of the privacy region. Figure 8 shows the 1st and 15th frames in the 'foreman' sequence, which are the I frame and P frame, respectively. Compared with the two corresponding frames in Figure 6, it is clear that the scrambling effect is effectively improved to conceal the details of the face. For example, in Figure 6, it is very easy to identity the position of eyes, the nose and the mouth. The emotions of the person in each frame can also be easily identified. In comparison, in Figure 8, all these details have been successfully scrambled, and therefore, the proposed technique provides better protection of the sensitive information. In Figure 9, the 15th frame of 'road', 'crew' and 'BQMall', and the 40th frame of 'hall', are also shown to verify the enhanced scrambling effect. In Figure 9(a), the skeleton of the car is still visible in the bottom-left corner, which could be utilized to determine its make and model, whereas, in Figure 9(b), this information is effectively scrambled. In Figure 9(c), the colour of the clothes can be easily observed, and even the man's belt can be seen. Once again in Figure 9(d), by using both IPM and SNC, this information can be effectively protected. In Figure 9(e), the facial regions of the two men in front are still visible: with the same effect as for 'foreman' in Figure 6, the relative position of eyes, the nose and the mouth in the facial regions can be easily recognized; and furthermore, the emotion of the left man is clearly visible. In comparison, in Figure 9(f), these facial regions are much more effectively scrambled to further conceal their details and emotions. Compared with Figure 9(g), the profile of the man on the left and the faces of the boy in the middle and the old woman on the right are scrambled more effectively in Figure 9(h). Thus, overall, it is very clear that compared with only using SNC, applying the combination of IPM and SNC can much more effectively enhance the scrambling effect for privacy region protection of videos with different resolutions.

It is also apparent from Figure 8 and Figure 9 that drift error is completely prevented by adopting the proposed re-encoding method and SWRME.

4.5. Savings in Bitrate Overhead for Privacy Region Protection

To evaluate the effect of the proposed re-encoding technique on the bitrate overhead, different quantization parameters (QP) are tested. The bitrate overhead is defined as the difference between the bitrate when privacy region protection is applied and the bitrate without privacy region protection. In Table 3, the bitrate overheads using the proposed re-encoding method, and the previous work by Tong *et al.* [23], which is the only other work that involves encryption of IPM and also addresses the incurred drift error, are compared. In Table 3, all frames are intra coded. It is clear that for different QP values, the proposed method can offer bitrate overhead savings over the previous work. In this paper, there is no direct comparison provided between the proposed work and the slice-based approaches to ROI encryption [13-16] [22]. As described, our work provides advantages compared with the previous work by Tong *et al.* [23] [24] and Tong's work has shown advantages compared with these slice-based approaches. For more details, please refer to [23] and [24]. The corresponding percentage of the bitrate overhead relative to the original bitrate without privacy region protection is also listed in Table 3. In Figure 10 and Figure 11, the intra period is set to different values in the baseline profile and the main profile, and the corresponding bitrate overhead savings are plotted. Here, the intra period is defined as the frame number difference of two adjacent I (Intra) frames, which decides the coding sequence. For example, in the main profile, intra period = 5 means that the coding sequence is 'IPBPBIPB...'. Generally as the intra period increases, the saving in bitrate overhead is decreased. This is because the proposed method only works for the intra prediction modes and the saving in bitrate overhead is highly dependent on the number of I macroblocks in the sequence.

5. Spiral binary mask to flag the position of the privacy region

This section discusses how to flag the position of the privacy region more efficiently. Generally, the privacy region in a video sequence can be retrieved by motion detection and object tracking as discussed in previous work [11] [13] [17]. In this work, it is assumed that the position of the privacy region is already known, and our focus is on how to efficiently '*represent*' the position.

Table 3. The bitrate overhead using the proposed method and the corresponding percentage relative to the original bitrate without privacy region protection, compared with the technique by Tong *et al.* [23] with all frames intra coded.

QP	Proposed Method										Tong *et al.* [23]									
	foreman		road		hall		crew		BQMall		foreman		road		hall		crew		BQMall	
	kb/s	%	kb/s	%	kb/s	%	kb/s	%	kb/s	%	kb/s	%	kb/s	%	kb/s	%	kb/s	%	kb/s	%
12	130.0	1.4	77.9	0.7	105.3	1.1	171.4	0.6	94.3	0.2	195.1	2.0	117.9	1.0	157.3	1.7	222.3	0.7	166.0	0.4
18	143.9	2.5	92.4	1.3	113.0	2.2	199.0	1.2	132.5	0.5	201.7	3.5	129.2	1.9	156.8	3.0	234.1	1.4	177.6	0.7
24	152.7	4.9	96.2	2.7	122.2	4.4	209.9	2.8	124.4	0.8	197.7	6.3	124.1	3.5	154.6	5.6	241.7	3.2	168.0	1.1
30	152.6	9.3	96.6	5.4	122.0	7.7	208.2	5.7	133.9	1.6	182.3	11.1	113.2	6.3	147.2	9.3	227.7	6.3	168.0	2.0
36	150.3	17.5	95.1	10.9	119.5	12.9	205.1	11.3	129.4	2.9	163.3	19.0	103.2	11.8	136.0	14.7	212.3	11.7	154.8	3.4

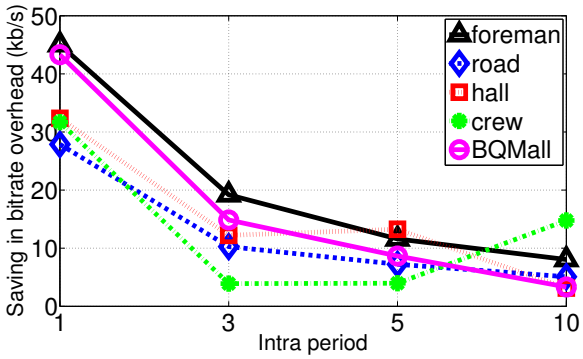


Figure 10. The bitrate overhead savings for different values of the intra period with QP=24 in the baseline profile.

5.1. The Binary Mask

Dufaux *et al.* [16] illustrated that the FMO technique can be adopted to indicate the position of the privacy region. Using FMO, intra prediction across slice boundaries is not permitted, thus it decreases the coding efficiency as indicated by Dai *et al.* [24]. To solve this problem, Dai *et al.* [24] proposed a binary mask (BM) mechanism to indicate the position of the privacy region, instead of using FMO. Their technique only incurs a relatively small bitrate overhead. The binary mask mechanism uses one bit per MB to signal whether or not it belongs to the privacy region, where a '1' and '0' indicate that it does or does not belong, respectively. The binary mask stream can be packeted into the SEI (Supplemental Enhancement Information) NAL (Network Abstraction Layer) and transmitted together with the video stream. The main intention of the FMO mechanism in H.264/AVC is to provide error resilience when slice loss occurs in an unreliable transmission channel. Both FMO and the binary mask mechanism adopt the same coding method

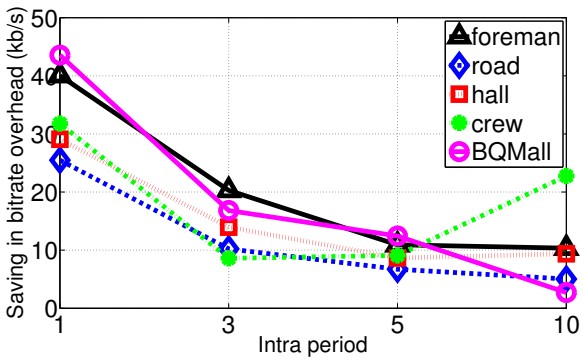


Figure 11. The bitrate overhead savings for different values of the intra period with QP=24 in the main profile.

to indicate the privacy region, one bit per MB. The difference between the two mechanisms is whether or not a frame can be encoded into more than one slice group, which will affect the coding efficiency and error resilience. The binary mask mechanism does not enable slice group coding as is used in FMO. Using the techniques explained in Section 3 and the binary mask mechanism, a significant improvement in the compression performance can be achieved [24]. However, as mentioned by Dai *et al.* [24], the binary mask itself still incurs a bitrate overhead of about 3Kb/s, 12Kb/s and 48Kb/s for a video at QCIF (Quarter CIF), CIF and 4CIF (4 × CIF) resolutions, respectively. This bitrate overhead can be further reduced using the proposed spiral binary mask described in the next section.

5.2. The Spiral Binary Mask

Generally, the region of interest in a video is defined as the face, or a moving object, and it only occupies a relatively low percentage of the overall frame [9 - 25]. Further more, there always exists a smallest rectangle in a frame to cover the privacy region.

In the example in Figure 12, the shaded region is a privacy region, denoted as P, and the corresponding non-privacy region is denoted as $\overline{P}$. The smallest rectangular region which can cover the privacy region is denoted as R which is highlighted by the bold dashed line in Figure 12. The bottom-left MB, 'c', and the two top-right MBs, 'd' and 'f', belong to $\overline{P} \cap R$.

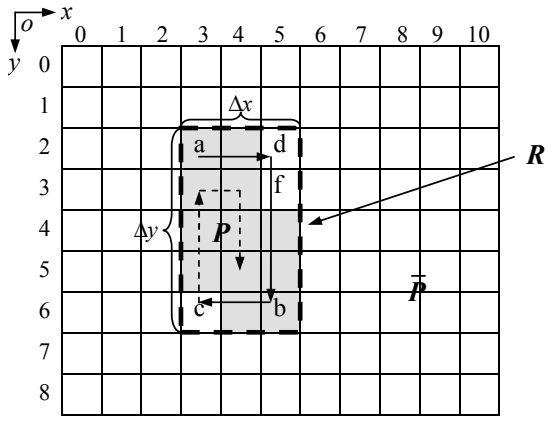


Figure 12. An example of a privacy region and the smallest rectangle in an 11x9 MBs frame(QCIF).

The rectangle R can be described by the top-left and bottom-right MBs, 'a' and 'b'. The coordinates of these two MBs are obtained according to the equations in (3).

$$\begin{cases} (x_a, y_a) = (\min x_i, \min y_i) \\ (x_b, y_b) = (\max x_i, \max y_i) \end{cases} \qquad here(x_i, y_i) \in P \tag{3}$$

The position of P can be represented by the coordinates of 'a' and 'b', and the binary mask of R. Since it is assumed that the coordinates of 'a' and 'b' have been obtained, the width Δx and height Δy of R can be calculated using the equations in (4). Each MB in R is assigned a one-bit binary mask, where '1' means that the corresponding MB belongs to P and vice versa. The binary mask is scanned one MB at a time in R. The bit length of the binary mask is equal to the number of MBs in R, which is the multiplication of Δx and Δy. From observation, most of the binary mask in R comprises consecutive bits with a value '1', and '0' appears sparsely. Thus, it is possible to further compress the bitstream of the binary mask. JBIG2 is an image compression standard for bi-level images, developed by the Joint Bi-level Image Experts Group (JBIG) [36], which can be used to compress this binary mask; however, the resulting relatively high complexity is not desirable and will affect the real time processing of video surveillance systems. To more efficiently represent the original binary mask, one technique would be to use two elements, the number of '0' bits that occur in the bitstream and the compressed binary mask obtained by using binary arithmetic

coding [31]. This method can effectively reduce the bit length to represent the binary mask in R but the computational complexity incurred by the binary arithmetic coding is relatively high. In this work, a spiral scan order is proposed, which takes advantages of the fact that the non-privacy MBs in R are very likely to appear on the four sides of R. As such the binary mask in R can be represented by the number of '0' bits in the bitstream, the spiral scan type and the spiral binary mask. Since the number of '0' bits has been explicitly signalled, the spiral scan will stop just before reaching the last '0' bit, and the resulting binary mask is referred to as a spiral binary mask. The last '0' bit is not included in the spiral binary mask. For example, as shown in Figure 12, the solid arrow line indicates the actual spiral scan order and the MBs under the dashed line will not be scanned. The spiral scan in Figure 12 starts from 'a' and stops before 'c' in a clockwise direction. The resulting spiral binary mask is '11001111'. The spiral scan order can start from any one of the four corners in R and progress in a clockwise or counter-clockwise direction as shown in Figure 13. There are eight types of spiral scan order. The spiral scan with the shortest spiral binary mask is chosen to represent the binary mask. To indicate which spiral scan type is used, a further three bits are needed: the first two bits indicate at which corner to start and the third bit indicates the direction.

$$\begin{cases} \Delta x = x_b - x_a + 1 \\ \Delta y = y_b - y_a + 1 \end{cases} \tag{4}$$

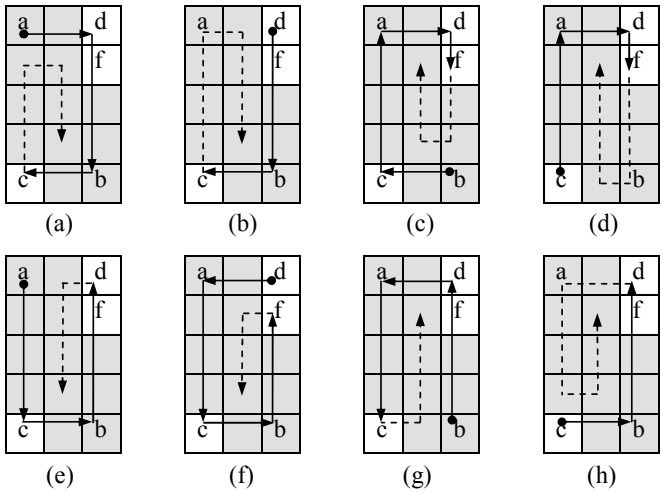


Figure 13. Illustration of the eight types of spiral scan: the scan starts from one of the four corners in R in a clockwise or counter-clockwise direction.

The syntax structure of the spiral binary mask is shown in Table 4. In order to effectively represent the coordinates of 'a' and 'b', two variables, *addr_a* and *addr_b*, are used and are calculated according to equation (5):

$$\begin{cases} addr\_a = y_a \times width\_frame + x_a \\ addr\_b = y_b \times width\_frame + x_b \end{cases} \tag{5}$$

In equation (5), $width\_frame$ is the number of MBs in a row in a frame. In Table 4, when the value of *number of zeros* is equal to 0, $R = P$, which means that the privacy region is rectangular. In this circumstance, *scan type* and *binary mask* will not appear in the syntax structure, because they will comprise all '1's. The bit lengths of *addr_a*, *addr_b* and *number of zeros* are denoted as len_a, len_b and len_{zero}, respectively. All of these are decided by the resolution of a frame. Further details for different frame resolutions are provided in Table 5. Here, l denotes the MB number in R, $l = \Delta x \times \Delta y$, and $\lceil \bullet \rceil$ represents the smallest integer not less than $\bullet$.

In addition, multiple spiral binary mask units using the proposed syntax definition in Table 4 can be concatenated and each of them can be decoded in sequence. Thus, it can be used to indicate multiple privacy

regions in a single frame.

Table 4. The syntax structure of the spiral binary mask

addr_a	*addr_b*	*number of zeros*	[*scan type*]	[*binary mask*]

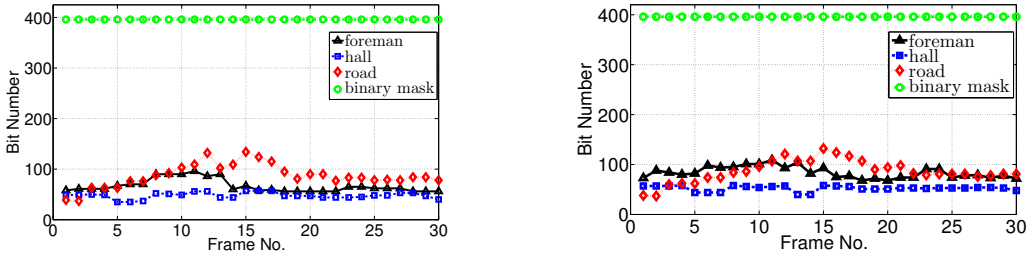
Table 5. The bit length of *addr_a*, *addr_b* and *zero number*

Resolution	MBs	len_a	len_b	len_{zero}
QCIF	11x9	7	7	$\lceil \log_2 l \rceil - 1$
CIF	22x18	9	9	
4CIF	44x36	11	11	

5.3. Savings in Bitrate Overhead to Flag the Privacy Region

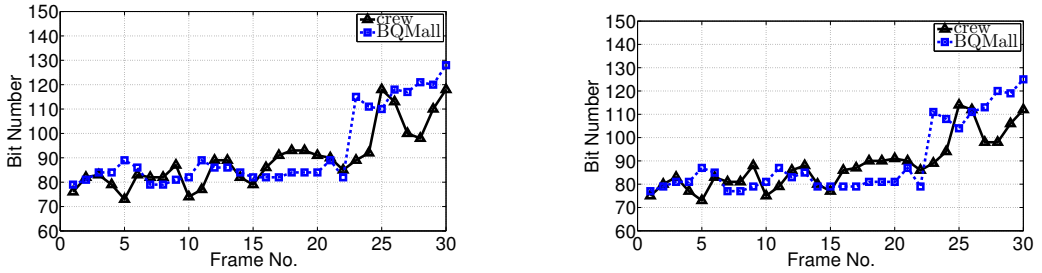
As indicated in Section 5, using the binary mask mechanism on a MB basis (one bit per MB) to flag the position of the privacy region will incur additional bitrate overhead. For a sequence at CIF resolution and 30 frames/s, each frame needs 396 bits, equivalent to about 12kb/s. For a sequence at 4CIF resolution and 30 frames/s, 1584 bits are required for each frame, which results in a bitrate overhead of about 48kb/s. The proposed spiral binary mask method can effectively reduce this overhead. Figure 14(a) shows the number of bits needed to indicate the privacy region using the proposed spiral binary mask for the three CIF sequences, 'foreman', 'road' and 'hall'. The equivalent bitrate overhead incurred by the spiral binary mask is about 1kb/s $\sim$ 3kb/s. Compared with the 12kb/s incurred using the binary mask mechanism for the CIF resolution, a 68% $\sim$ 90% reduction is achieved. In the sequence, 'road', there are multiple privacy regions, such as the 3 cars in the 15th frame as shown in Figure 9(a). For each of the privacy regions, about 20 bits are needed to indicate the two corners of the smallest rectangle and the number of zero bits inside it, which are the first 3 syntax elements in Table 4. Thus, the performance for 'road' is degraded to some extent, when compared with the other two test sequences. The results for the two 4CIF sequences, 'crew' and 'BQMall', are shown in Figure 15(a). Compared with the 48kb/s incurred using the binary mask mechanism for 4CIF resolution, there is only a bitrate overhead of less than 3kb/s when the proposed spiral binary mask mechanism is employed, which means a much greater saving (more than 93 %) is obtained. This result further verifies the effectiveness of the proposed spiral binary mask mechanism. Comparing the results in Figure 14 and 15, for different resolutions, the bit number of the spiral binary mask to indicate the privacy region in each frame does not increase the bitrate too much, while for the binary mask, the increase of the resolution from CIF to 4CIF results in four times the bitrate overhead. Thus, a greater saving percentage is observed at a higher resolution.

Previously, the authors proposed to encode the binary mask of the smallest rectangle for each privacy region using the binary arithmetic coding [31]; however the binary arithmetic coding method was relatively complicated compared with the spiral binary mask proposed here. In comparison, the results using the authors' previous work are shown in Figure 14(b) and Figure 15(b). It can be seen that the spiral binary mask can achieve the same effectiveness as the previous work but without using complex arithmetic coding. To give a more explicit comparison, the bitrate overhead is summarized in Table 6. The proposed spiral binary mask can achieve a better result for 'foreman' and 'hall" and very slightly worse for the other three sequences, as shown in Table 6. The binary arithmetic coding exploits the probabilistic distribution of the binary mask in the smallest rectangle covering each privacy region. The spiral binary mask benefits from the side information that most of the '0' bits (corresponding to non-privacy macroblocks in the smallest rectangle) are around the perimeter of the smallest rectangle. This side information can contribute much higher bitrate savings when the privacy region is larger. In 'foreman' and 'hall', the privacy region is slightly larger than the other three as shown in Table 6.



(a) Spiral binary mask (b) Binary mask coded by arithmetic coding

Figure 14. The bit number at each frame to indicate the privacy region in 'foreman', 'hall' and 'road'.



(a) Spiral binary mask (b) Binary mask coded by arithmetic coding

Figure 15. The bit number at each frame to indicate the privacy region in 'crew' and 'BQMall'.

6. Conclusions and future work

In the majority of previous work on privacy region protection, only the sign bits of nonzero coefficients (SNC) in the privacy region are proposed to be encrypted. However, the scrambling effect of SNC is relatively weak. Thus, it is proposed to also encrypt the intra prediction modes (IPM) together with SNC in the privacy region, which can effectively enhance the scrambling effect. Directly encrypting the intra prediction modes will incur drift error in the non-privacy region. In this paper, a re-encoding method is proposed to remove the drift error caused by applying IPM in the privacy region. The intra prediction modes in the non-privacy region adjacent to the privacy region are re-encoded based on the encrypted IPMs in the privacy region. In this case, when decoded without decryption, the intra prediction modes in the non-privacy region can be decoded correctly and no drift error will occur. Compared with the technique reported by Tong *et al.* [23] which also uses encryption of IPM, the proposed method offers savings in the bitrate overhead. Experimental results and analysis based on H.264/AVC for CIF and 4CIF resolutions were carried out and verified the effectiveness of the proposed methods.

A spiral binary mask mechanism is also proposed to reduce the bitrate overhead incurred by indicating the position of the privacy region, which is transmitted as additional information. Using the proposed

Table 6. The bitrate overhead to indicate the privacy region by using the proposal spiral binary mask (denoted as 'Spiral BM'), the authors' previous work (denoted as 'Arithmetic BM') [31], and the binary mask (denoted as 'BM'). Unit: bit/s

	CIF			4CIF	
	foreman	hall	road	crew	BQMall
Spiral BM	1998	1434	2605	2684	2778
Arithmetic BM	2521	1572	2597	2649	2695
BM	11880			47520	46800

spiral binary mask, a bitrate overhead of approximately 1kb/s $\sim$ 3kb/s is incurred at CIF resolution and 30 frames/s. Compared with the bitrate overhead of approximately 12kb/s incurred by the binary mask method [23] [24], a 68% $\sim$ 90% reduction is achieved. At 4CIF resolution, a much great saving of over 93% in the bitrate overhead is observed.

Overall, utilising the techniques proposed in this paper, the privacy regions in video sequences can be effectively protected with an enhanced scrambling effect without drift error and with a very low bitrate overhead.

Future work will include investigating privacy region protection in the latest video compression standard, HEVC. Another interesting topic would be to investigate the dependency between the effectiveness of the existing video encryption methods and the mode decision process of the video compression. There is limited research on possible attacks of the existing video encryption and privacy region protection techniques, which is also a further direction for possible future research.

References

[1] A. Senior, S. Pankanti, A. Hampapur, L. Brown, Y. Tian, A. Ekin, J. Connell, C. Shu, M. Lu, Enabling Video Privacy through Computer Vision, IEEE Security & Privacy 3 (3) (2005) 50–57.

[2] M. L. Gras, The Legal Regulation of CCTV in Europe, Surveillance & Society 2 (2/3) (2004) 216–229.

[3] A. Cavallaro, Privacy in Video Surveillance, IEEE Signal Process. Mag. 24 (2) (2007) 166–168.

[4] A. Martínez-Ballesté, P. A. Pérez-Martínez, A. Solanas, The Pursuit of Citizens' Privacy: A Privacy-Aware Smart City Is Possible, IEEE Communications Mag. 51 (6) (2013) 136–141.

[5] F. Porikli, F. Bremond, S. Dockstader, J. Ferryman, A. Hoogs, B. Lovell, S. Pankanti, B. Rinner, P. Tu, P. Venetianer, Video Surveillance: Past, Present, and Now the Future [DSP Forum], IEEE Signal Process. Mag. 30 (3) (2013) 190–198.

[6] H.264, Advanced Video Coding for Generic Audio-visual Services, ISO/IEC 14496-10:2010 ITU-T REC.

[7] I. Richardson, The H. 264 Advanced Video Compression Standard, Wiley, 2010.

[8] Y. Shi, H. Sun, Image and Video Compression for Multimedia Engineering: Fundamentals, Algorithms, and Standards, CRC Pr I Llc, 2008.

[9] E. Newton, L. Sweeney, B. Malin, Preserving Privacy by De-identifying Face Images, IEEE Trans. Knowledge & Data Engineering 17 (2) (2005) 232–243.

[10] W. Zhang, S. Cheung, M. Chen, Hiding Privacy Information in Video Surveillance System, in: Proc. 12th IEEE Int. Conf. Image Process. (ICIP), 2005, pp. 868–871.

[11] I. Martínez-Ponte, X. Desurmont, J. Meessen, J. Delaigle, Robust Human Face Hiding Ensuring Privacy, in: Proc. Int. Workshop Image Analysis for Multimedia Interactive Services (WIAMIS), 2005.

[12] X. Yu, K. Chinomi, T. Koshimizu, N. Nitta, Y. Ito, N. Babaguchi, Privacy Protecting Visual Processing for Secure Video Surveillance, in: Proc. 15th IEEE Int. Conf. Image Process. (ICIP), 2008, pp. 1672–1675.

[13] F. Dufaux, M. Ouaret, Y. Abdeljaoued, A. Navarro, F. Vergnenègre, T. Ebrahimi, et al., Privacy Enabling Technology for Video Surveillance, in: Proceedings of SPIE, Vol. 6250, 2006, pp. 205–216.

[14] F. Dufaux, T. Ebrahimi, Scrambling for Video Surveillance with Privacy, in: Proc. IEEE Workshop Computer Vision & Pattern Recognition, 2006, pp. 160–166.

[15] F. Dufaux, T. Ebrahimi, Scrambling for Privacy Protection in Video Surveillance Systems, IEEE Trans. Circuits Syst. Video Technol. 18 (8) (2008) 1168–1174.

[16] F. Dufaux, T. Ebrahimi, H.264/AVC Video Scrambling for Privacy Protection, in: Proc. 15th IEEE Int.Conf. Image Process. (ICIP), 2008, pp. 1688–1691.

[17] N. Baaziz, N. Lolo, O. Padilla, F. Petngang, Security and Privacy Protection for Automated Video Surveillance, in: Proc. IEEE Int. Symp. Signal Process. & Info. Technol., 2007, pp. 17–22.

[18] P. Carrillo, H. Kalva, S. Magliveras, Compression Independent Object Encryption for Ensuring Privacy in Video Surveillance, in: Proc. IEEE Int. Conf. Multimedia & Expo (ICME), 2008, pp. 273–276.

[19] K. Martin, K. Plataniotis, Privacy Protected Surveillance Using Secure Visual Object Coding, IEEE Trans. Circuits Syst. Video Technol. 18 (8) (2008) 1152–1162.

[20] H. Sohn, E. AnzaKu, W. De Neve, Y. Ro, K. Plataniotis, Privacy Protection in Video Surveillance Systems Using Scalable Video Coding, in: Proc. 6th IEEE Int. Conf. Advanced Video & Signal Based Surveillance (AVSS), 2009, pp. 424–429.

[21] Y. Wang, M. O'Neill, F. Kurugollu, Privacy Region protection for H.264/AVC by Encrypting the Intra Prediction Modes without Drift Error in I Frames, in: Proc. IEEE Int.Conf. Acoustic, Speech & Signal Process. (ICASSP), 2013, pp. 2964–2968.

[22] F. Peng, X.-w. Zhu, M. Long, A ROI Privacy Protection Scheme for H. 264 Video Based on FMO and Chaos, IEEE Trans. Info. Forensics & Security (2013) 1–12.

[23] L. Tong, F. Dai, Y. Zhang, J. Li, Restricted H. 264/AVC Video Coding for Privacy Region Scrambling, in: Proc. 17th IEEE Int. Conf. Image Process. (ICIP), 2010, pp. 2089–2092.

[24] F. Dai, L. Tong, Y. Zhang, J. Li, Restricted H. 264/AVC Video Coding for Privacy Protected Video Scrambling, J. Visual Communication & Image Representation 22 (6) (2011) 479–490.

[25] S. Choi, G. Kim, J. Han, On the Challenges of Applying Selective Encryption on Region-of-Interest in H. 264 Video Coding, in: Proc. 15th IEEE Int. Conf. Computer Science & Its Applications, 2009, pp. 1–5.
[26] C. Shi, B. Bhargava, An Efficient MPEG Video Encryption Algorithm, in: Proc. 17th IEEE Symp. Reliable Distributed Systems, 1998, pp. 381–386.
[27] B. Bhargava, C. Shi, S. Wang, MPEG Video Encryption Algorithms, Multimedia Tools & Applications 24 (1) (2004) 57–79.
[28] J. Ahn, H. Shim, B. Jeon, I. Choi, Digital Video Scrambling Method Using Intra Prediction Mode, Advances in Multimedia Information Processing-PCM 2004 (2005) 386–393.
[29] S. Lian, Z. Liu, Z. Ren, H. Wang, Secure Advanced Video Coding Based on Selective Encryption Algorithms, IEEE Trans. Consumer Electron. 52 (2) (2006) 621–629.
[30] Y. Wang, M. O'Neill, F. Kurugollu, A Tunable Encryption Scheme and Analysis of Fast Selective Encryption for CAVLC and CABAC in H.264/AVC, Accepted by IEEE Trans. Circuits Syst. Video Technol. (2013) 1–15.
[31] Y. Wang, M. O'Neill, F. Kurugollu, Adaptive Binary Mask for Privacy Region Protection, in: Proc. IEEE Int. Symp. Circuits & Systems (ISCAS), 2012, pp. 1127–1130.
[32] P. Lambert, W. De Neve, Y. Dhondt, R. Van de Walle, Flexible Macroblock Ordering in H. 264/AVC, J. Visual Communication & Image Representation 17 (2) (2006) 358–375.
[33] S. Lian, Z. Liu, Z. Ren, Z. Wang, Selective Video Encryption Based on Advanced Video Coding, Advances in Multimedia Information Processing-PCM 2005 (2005) 281–290.
[34] JVT, [online] http://iphome.hhi.de/suehring/tml/.
[35] M. Boesgaard, M. Vesterager, T. Christensen, E. Zenner, The Stream Cipher Rabbit, ECRYPT Stream Cipher Project Report 6.
[36] P. G. Howard, F. Kossentini, B. Martins, S. Forchhammer, W. J. Rucklidge, The Emerging JBIG2 Standard, IEEE Trans. Circuits Syst. Video Technol. 8 (7) (1998) 838–848.